How Can Predictive Learning Analytics and Motivational Interventions Increase Student Retention and Enhance Administrative Support in Distance Education?

Journal Item

JOURNAL OF LEARNING ANALYTICS

SoLAR SOCIETY for LEARNING ANALYTICS RESEARCH

Volume 7(2), 72-83. http://dx.doi.org/10.18608/jla.2020.72.4

How Can Predictive Learning Analytics and Motivational Interventions Increase Student Retention and Enhance Administrative Support in Distance Education?

Christothea Herodotou[1], Galina Naydenova[2], Avi Boroowa[3], Alison Gilmour[4], Bart Rienties[5]

Abstract

Despite the potential of Predictive Learning Analytics (PLAs) to identify students at risk of failing their studies, research demonstrating effective application of PLAs to higher education is relatively limited. The aims of this study are 1) to identify whether and how PLAs can inform the design of motivational interventions and 2) to capture the impact of those interventions on student retention at the Open University UK. A predictive model — the Student Probabilities Model (SPM) — was used to predict the likelihood of a student remaining in a course at the next milestone and eventually completing it. Undergraduate students (N=630) with a low probability of completing their studies were randomly allocated into the control (n=312) and intervention groups (n=318), and contacted by the university Student Support Teams (SSTs) using a set of motivational interventions such as text, phone, and email. The results of the randomized control trial showed statistically significant better student retention outcomes for the intervention group, with the proposed intervention deemed effective in facilitating course completion. The intervention also improved the administration of student support at scale and low cost.

Notes for Research

- This paper explains how predictive learning analytics can help with the selection of students who need support.
- A motivational intervention delivered via text, phone, and email was tested.
- Students receiving the intervention had statistically significant better retention outcomes compared to the control group.
- Student Support Teams could use predictive learning analytics to target specific students and better support them.

Keywords

Predictive learning analytics, motivational interventions, student support, distance learning, randomized control trial

Submitted: 17.07.2019 — **Accepted:** 15.05.2020 — **Published:** 19.09.2020

Corresponding author [1] Email: christothea.herodotou@open.ac.uk Address: The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, UK. https://orcid.org/0000-0003-0980-1632

[2]Email: galina.naydenova@open.ac.uk Address: The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, UK

[3]Email: avi.boroowa@open.ac.uk Address: The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, UK

[4]Email: alison.gilmour@open.ac.uk Address: The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, UK

[5]Email: bart.rienties@open.ac.uk Address: The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, UK

1. Introduction

Predictive Learning Analytics (PLAs) have the potential to help institutions identify which groups and individual students, often described as "at risk," might need extra support to reach desired learning outcomes (Conijn, Snijders, Kleingeld, & Matzat, 2017; Kovanović et al., 2015; Scheffel et al., 2017; Tempelaar, Rienties, Mittelmeier, & Nguyen, 2018). Yet, outcomes around their effectiveness remain blurred. For example, Scheffel et al. (2017) found that visualizing the relative performance of 172 students — by particular initiative (number of posts made) and responsiveness (number of comments to posts) within a group — supported learning processes over time, whereas presence (number of page views) was found to be negatively related to student performance. Although institutions are interested in learning analytics, their actual uptake and integration in most institutions around the world is rather limited (Ferguson et al., 2016; Sclater, Peasgood, & Mullan, 2016; West, Luzeckyj, Searle, Toohey, & Price, 2018). Available evidence draws from relatively small case studies of PLAs usage with teachers (van Leeuwen, Janssen, Erkens, & Brekelmans, 2015), experimental lab studies (Harley et al., 2016; Worsley & Blikstein, 2015), and MOOC contexts (Kizilcec, Pérez-Sanagustín, & Maldonado, 2017) and shows that PLAs can have a positive effect on learners and learning outcomes. The few large-scale studies available also point to positive outcomes when teachers engage systematically with PLAs (Herodotou et al., 2020; 2019a; 2019b).

Although there is promising initial support of PLAs to support learning, there is limited application and evaluation of them to the design of motivational interventions and consideration of what this means in terms of institutional approaches to supporting retention. An example study is a retention campaign program designed to advise students at risk about available support options, with the intentions of enhancing student resilience and increasing student success and retention. The program evaluation showed no effects on retention outcomes; yet individual differences were more likely to explain retention patterns (Dawson, Jovanovic, Gašević, & Pardo, 2017; Ferguson & Clow, 2017; Sclater, 2017; Seidel & Kutieleh, 2017).

In this study, we designed, delivered, and evaluated a relatively large motivational intervention with 630 students identified through PLAs as being at risk of not completing their studies. Student Support Teams (SSTs) at the Open University UK delivered the proposed intervention and a randomized control trial (RCT) was used to test its effectiveness. The aim of the study was to showcase how PLAs can inform the practice of SSTs in identifying which motivational intervention can effectively reinforce student engagement and course completion.

1.1. Predictive Learning Analytics and Motivational Interventions

PLAs are defined as analytics "forecasting future outcomes and behaviours […] the forecasts are achieved by extracting and processing information from data routinely collected during a business process" (Calvert, 2014, p. 161). A number of PLAs used in conjunction with early alert systems have been developed over the years aiming at the early identification of students at risk and the provision of timely interventions, such as Course Signals (Arnold & Pistilli, 2012), E2 Coach (Wright, McKay, Hershock, Miller, & Tritz, 2014), and OU Analyse (Hlosta, Zdrahal, & Zendulka, 2017). PLAs have allowed higher education institutions to reconsider the potential of data on a different scale than previously, through utilizing and acting upon data currently captured within university systems. Arnold and Pistilli (2012) argued that an essential element of effective retention policies is to support the "academic integration" of students; the use of learning analytics incorporating both demographic and behavioural data can support student persistence. The application of learning analytics data produces "actionable intelligence," whereby identifying students most at risk of not continuing their studies should allow for early intervention by the institution (p. 267). PLAs have to be accompanied by effective motivational interventions if they are to influence student outcomes (Jayaprakash, Moody, Lauría, Regan, & Baron, 2014).

Within distance learning in particular, the challenge of identifying students experiencing difficulty and offering support is perhaps even more prominent than at campus-based universities, where issues may manifest more evidently in participation in classroom activities, potentially leading to earlier detection. Simpson (2013) proposed the "distance education deficit" (p. 105) to highlight the challenges faced in distance education in terms of course retention and qualification achievement. Key to this work was the exploration of factors impacting retention, including the identification of learner motivation as a significant determining factor. Moore's (1997) notion of "transactional distance" is of relevance here; it signposts the significance of both the physical and psychological space between students and their teachers, and of communication as one element influencing transactional distance in distance education. In the context of retention, narrowing the gap or transactional distance between student and institution, potentially through communication, and minimizing the opportunities for students to exit prematurely from their studies, are both areas of focus within this retention study.

Different interventions can take place and influence student outcomes. Motivational interventions (e.g., Wentzel & Wigfield, 2007) aim to increase student engagement and retention using various mediums, from direct phone calls to emails and text messages. On the other hand, pedagogical interventions (e.g., Wise, 2014) aim to enhance student learning using methods such as adaptive learning platforms, online tutoring, the provision of additional, tailored study resources, and problem solving exercises via the motivational mediums described above. In this study, we address the first type of intervention — motivational — with the aim of reinforcing student participation and increasing retention. The university under study has an "open entry policy," meaning that students can register without requiring such prerequisites as a high school certificate. As opposed to filtered entry, open entry has much higher student withdrawal rates (van Ameijde, Weller, & Cross, 2018), hence the focus of this study on testing an intervention that may influence student attrition.

Various motivational interventions have been tested over the years at the university under study. For example, the use of texts as a means of administrative communication with students about such things as room changes and assignment submission procedures was positively endorsed by students (Naismith, 2007). Other studies examined the composition of emails and how they can be a motivator for effective, efficient, engaging learning (Kim, 2008); mobile learning (Fozdar & Kumar, 2007); interactive online tools that enable students to access "the right information, at the right time, in the right place" (Ryan & Davies, 2016); and robot teaching assistants for improving performance and intention to continue studying (Hung, Chao, Lee, & Chen, 2013). At distance learning institutions, what is often tested are approaches delivered by SSTs — teams actively supporting and monitoring student engagement and progress over time (Ryan & Davies, 2016; Simpson, 2013). While proactive support by SSTs can be valuable in reaching students who face difficulties, this approach is rather resource intensive for an institution, possibly resulting in selecting only a narrow group of students at risk. In this respect, PLAs could be a useful tool for targeting motivational interventions towards students in most need of support, or focusing on students for whom the intervention could have the largest impact. Research shows that elements of PLAs are gradually added as amendments or "layering" to existing SST approaches (Dolata, 2013), and as a means to overcome the lack of financial and human resources.

Few studies presenting mixed findings have examined the impact of motivational interventions on student retention. For example, using a predictive model, Dawson et al. (2017) identified first-year university students at risk of academic success and retention and piloted a retention program consisting of calling students and providing advice and support. While both Chi-square and logistic regression analyses revealed a significant association between the intervention and better retention outcomes, a follow-up logistic mixed-effects analysis showed that improvements in retention were explained by student characteristics, not by the intervention itself. In contrast, Castleman and Page (2015) tested the impact of a series of low-cost nudges and motivational interventions on college enrollment following graduation. These interventions consisted of automated text messages of tasks to complete the form for entering college and requesting help when needed, as well as peer mentors proactively reaching out students and offering college transition support. After allocating students into control and intervention groups, a significant increase in the number of student enrollments were observed, especially within populations often underrepresented in Higher Education. Through an RCT, Inkelaar and Simpson (2015) tested the retention impact of sending motivational emails to students every two weeks and identified a significant increase of 2.3% in retention ($p<.10$) in the intervention group. Despite the low increase, the authors indicated that the costs of the intervention (i.e., £1,000 for using a system of motivational emails) is substantially lower than the increased revenue from 2% more final exam registrations (i.e., £7,500). They also noted the importance of proactive motivational support and the composition of emails that promote interactivity, nudging, and priming.

In other words, while there is some small-scale evidence that intervention using PLAs in combination with support actions (e.g., emails, text messages) can have a small, positive effect, there is a paucity of research at the very early stages of enrollment at universities. As highlighted by Castleman and Page (2015) and Shea and Bidjerano (2014), the early stages of enrollment and study can be particularly stressful for some groups of students, but at the same time essential for gaining an early advance to a successful start. Therefore, in this study we explored whether, six weeks before the start of a course, an introductory text message and a follow-up phone call or email with students potentially at risk (as identified by PLAs) could have an immediate, longer-term impact on student retention.

2. Methods

At the distance university under study, student support is organized within a curriculum-aligned SST model. SSTs utilize a framework for proactive student support known as Model of Integrated Learning and Learner Support (MILLS; Beard, Stevens, & Welch, 2017). This framework defines how the principles and practices within each SST should function in order to support students effectively. Within MILLS, the suggested interventions are a form of proactive student support and can take different forms. Contact can be via automated emails, direct emails, or phone calls from either the student's tutor or the SSTs at designated points within the MILLS framework. One example is a follow-up on the non-submission of an assignment, or an email sent to students about getting ready for the final exam. Interventions are applied when students enroll in their first course and during the course implementation. SSTs contact students outside normal office hours (given that most students are part-time and committed to other activities during the day). A 7% increase in student retention since proactive student support was introduced was recorded, as well as a 3% improvement in course completion when students were contacted by email. Proactive phone calls were also found to motivate students and boost their confidence (Wilcock, 2011).

In this study, the MILLS intervention under examination, coded as M2 ("Study preparation"), aimed to ensure that at-risk students were actively contacted and were on the right track with their study plans. Contact took place before the start of the course and after enrollment. The motivational intervention included either an email or a proactive phone call to all at-risk students and covered the following areas: introducing the SSTs, exploring issues of motivation, interest, and employability, and preparing to study. Contact was also used to discuss workload, study intensity, and module choice, encouraging students to contact their teachers and/or SSTs if they faced any study issues that could not be resolved by online resources and information. The M2 intervention acted to reduce the "transactional distance" between the institution and the student, and was clearly informed by Simpson's (2013) work stressing the importance of learner motivation in retention. To identify the cohort of students receiving the M2 intervention, the following indicators, or a selection of them, were used by the SST under study: no formal qualifications, insufficient academic progress (IAP) alert or restricted,[1] registered with more than 120 credit courses, and low Index of Multiple Deprivation (IMD) (bottom quintile). Each faculty advised SSTs as to which indicators should be used for selecting students based on, for example, awareness of the level of difficulty of certain courses and their potential impact on student workload and completion. This approach was more likely to have had an impact on the outcomes of the M2 intervention in terms of correctly selecting the students in most need of support.

To overcome possible bias in student selection, in this study, SSTs identified students at risk of early exit from their studies (hereafter referred to as "at risk") using PLAs, in particular a statistical model that identified students at risk based on data related to continuous and final assessment, study events, student profile, study progression, virtual learning environment (VLE) engagement and other factors. Those students received a set intervention comprised of an initial text informing them that they would be contacted shortly by phone, followed by a phone call. Students who could not be reached by phone received an email.

2.1. Student Probabilities Model (SPM)

Table 1. SPM Variables Predicting Success and Retention (adapted from Calvert, 2014)

Student	**Student: Previous studies/reasons for study**
1. Price area: Student loan situation varies by country	6. Previous studies/reasons for study
2. Segment: Several sociodemographic factors	7. Credit transfer
3. Index of multiple deprivation	8. Highest qualification code on entry
4. Occupational status	9. Study motivation code
5. Disability flag	10. Study goal code
Student progress in previous studies at the same university	11. Highest award intention 12. Sponsored
13. OU credits already achieved	**Student: Module**
14. Still needed to total 360 credits	19. Total number modules registered
15. Number of "false starts"	20. Total number of credits studying in year
16. Student's total number of previous fails	21. Flag for over 120 credits at registration
17. Student's total number of previous withdrawals	22. Proportion of days through module
18. Student's total number of previous passes	23. Proportion of assignments completed at milestone dates
Qualification/module of study	24. Latest assignment score on module
26. Module	25. Earliest assignment score on module
27. Length of presentation of module	
28. Number of course assignments at end/milestone	
29. Intensity: Number of course assignments per day of module	
30. Qualification code	

The Student Probabilities Model (SPM) produced predictions of whether an individual student would reach specific milestones (different points in a course presentation or between courses), such as completing and passing a course, or returning in the next academic year. The SPM was initially developed by Calvert (2014), and further fine-tuned over time through a number of studies in which the probabilities were used under experimental conditions. SPM predictions or probabilities were based on models generated through logistic regression of a set of 30 explanatory variables (see Table 1).[2] These variables were grouped in student factors (IMD area, price area, disability, etc.), students' previous study (highest qualification on entry, etc.),

[1] IAP is a regulation of students who fail to make academic progress on modules at undergraduate level; if in three eligible presentations of study, a student does not complete at least one course successfully, they will not be allowed to register for another module unless they meet the conditions specified by the Senate.

[2] SPM has been recently redesigned considering for additional predictive factors such as previous higher education experience and VLE engagement in terms of number of visits averaged over the most recent period between the course milestones.

students on a course (total credits studying in a year, late registration, etc.), students' previous progress at the university (best previous score, number of fails, etc.), and course and qualification variables. These variables were chosen based on the availability of data in existing datasets and due to their relationship to success in a given course presentation and/or students returning to their studies the next academic year. A different set of factors will predict each course milestone. In this study, SPM was used by SSTs to identify students at a low probability of completing their studies so that they could target available resources towards that group of students. Ours is the first large-scale study using a comparative experimental methodology (RCT).

2.2. Study Design

We first compared the existing SST method of selecting students at risk of not completing their studies to using the SPM as an alternative method of selection. The following indicators were used by the SST: no formal qualifications, insufficient academic progress (IAP) alert or restricted, registered with more than 120 credit courses, and low IMD (bottom quintile). Several students were identified per indicator as potentially not completing their studies. This number was then compared to the actual number of students completing their courses. Table 2 compares the number of students selected per indicator to the students selected through the SPM as well as the percentage of students correctly identified as at risk, as measured by the actual number of students completing their courses. Overall, SPM was shown to correctly identify a greater percentage of students at risk of not completing their studies (55%) compared to such indicators as qualifications, credits, and IMD. A follow-up analysis with a different cohort, using the revised version of SPM showed even stronger outcomes; 73% of STEM students were correctly identified as at risk by SPM compared to 49–57% using other indicators. Student probabilities enabled the successful identification of a higher number of potential non-completers than the use of a single indicator. The SPM is a more reliable method of identifying at-risk students, early in the lifecycle of a course or before the course presentation starts, than existing single indicators.

Table 2. Comparing Different Methods of SST Selection of Students At Risk

Selection method	Number of students selected as at risk (predicted)	Number of students not completing their courses (actual)	% of selection that did not complete (actual)
Student Probabilities	1,968	1,088	**55%**
Indicator 1: No formal qualifications	295	124	42%
Indicator 2: IAP alert or restricted	459	252	55%
Indicator 3: More than 120 credits	132	45	34%
Indicator 4: Low IMD (bottom quintile)	1,795	738	41%

Prior to the SPM introduction, the standard SST practice was to get in touch with as many at-risk students as possible based on available resources. The choice of students was rather random, based on the order of the list of students at risk, with students at the top of the list more likely to be contacted by SSTs. To address the challenge of limited resources, we used SPM data to narrow the selection range, refining students in probability bands (see Table 3) and identifying a band that suited the capacity and resources available (i.e., how many students could be contacted by SSTs at the given point of time). In this study, the selected probability band was 31–40%.

Table 3. Using SPM to Split Students At Risk into Probability Bands

Probability band	Number of potential students at risk of not completing	% of students not completing their studies
0–10%	44	87%
11–20%	184	68%
21–30%	305	66%
31–40%*	**513**	**54%**
41–50%	922	48%

** Students who took part in this study.*

SPM data were generated when students from a mixture of undergraduate courses and levels across the university reached registered status. Participating students came from a university in Scotland, hence the reference to the "Scottish cohort" in Figure 1. Students with a 31–40% probability of completing their course were selected for this study. This was a pragmatic choice based on the availability of staff to implement the intervention. Furthermore, given the relatively small intervention treatment, the impact on students below 30% probability was not expected to be sufficient to help them, while students above 40% probability might have sufficient resilience to continue. To design the study (see Figure 1), we first used the SPM to identify those students with a 31–40% probability of completing their course for inclusion in the study (N=410). The second step was to randomize the students into two groups, 200 in the control group and 210 in the intervention group (see HC: Head Count). Note that some students were registered in more than one course, including courses where their probability of completion was other than 31–40%. The Student Module Count (SMC) indicates the number of students registered in more than one course. The analysis in Figure 1 took into account the SMC. All participating students (control and intervention) received the M2 intervention. The intervention group received an additional text from SSTs followed by a phone call or email (for those who did not answer the phone).

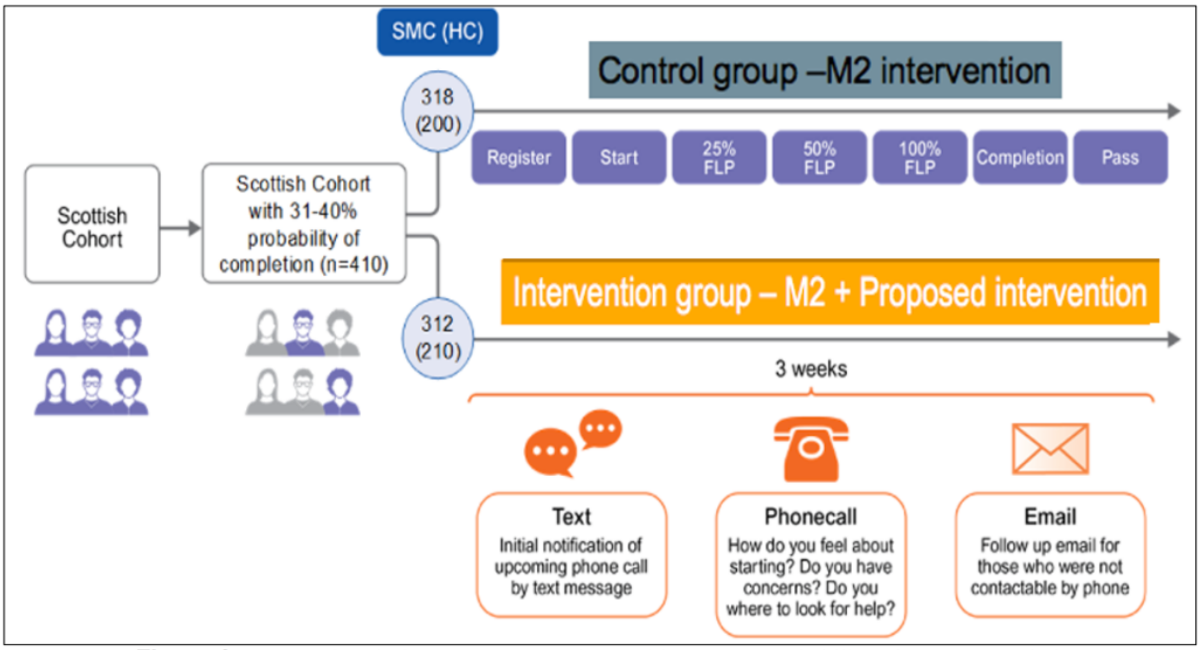

Figure 1. Intervention design to support students with 31–40% probability of course completion.

Both the control and intervention groups had similar chances of receiving the M2 intervention by SSTs. After inspecting the randomized data, only three students in the control group and seven in the intervention group had actually received the M2 intervention, indicating that the standard approach of selecting students for the M2 intervention was rather different to the SPM. In addition to M2, the intervention group received an initial text saying they would be contacted shortly by phone (n=210), followed by the phone call (n=207, of which n=91 calls reached the students). Students who could not be reached by phone received an email (n=119) (see Table 4). Proactive calls and emails were utilized because evidence pointed to their effectiveness in improving retention in distance education (e.g., Wilcock, 2011). A text was deemed appropriate to prevent a "sudden call" from an unknown number.

Table 4. Motivational Intervention Design

Contact activity	N=	%
Text sent	210	100%
Call attempted	207	98.6%
Call successful	91	43.3%
Email sent	119	56.7%

Texting had also been used in the past by SSTs as a useful way of reaching students. During the phone call, students were asked the following:

1. How do you feel about starting your studies/next course?
2. Do you have any questions about getting started?
3. Do you feel you have the information you need to begin your studies?
4. Do you have any additional questions?
5. Has this call been helpful to you today?

Specific notes were provided on a spreadsheet to assist SSTs with the discussion.

Drawing from previous experience, SSTs advocated that contact with students should take place as early as possible for greater impact. Therefore, the intervention took place after registration and six weeks before the start of the courses. The control and intervention groups were compared at different course milestones structured around the university's fee liability points (FLPs; i.e., dates when students who wish to discontinue their studies can claim fee credits or a refund): 1) percentage of student registrations at the start of a course (see Start, Figure 1); 2) registrations at the 25% mark (25% FLP; 14 days after the course start); 3) 50% (50% FLP; 3 months after course start); 4) 100% (100% FLP; 6 months after course start); and 5) course completion (end of the course). In terms of potential sampling bias, using independent t-tests, we checked for any differences in the control and intervention groups in terms of gender, age, prior educational qualifications, and socioeconomic background; we found no significant differences.

3. Results and Discussion

To address the research question of this study, we initially compared the retention rates at different milestones in the control and intervention groups as a whole. We then ran a more fine-grained analysis comparing the two different types of interventions (phone call versus email) with the control group, to identify whether one intervention was more effective than the other. Chi-square analyses were performed to compare the control and intervention groups at the four course milestones where a student might exit the course (see Table 5). Given the multiple comparisons, the p-value (.05) was adjusted accordingly (p=.05/5=.01). Strong statistically significant differences were observed at all milestones, suggesting that the proposed intervention as a whole (texting, phone call, email) was successful, and likely helped students remain engaged and progress through their studies. As shown in Figure 2, the proportion of students in the intervention group, compared to the control group, is increasingly larger over the lifecycle of participating courses, starting with 7.4% more students at the start of the courses, and reaching 22.2% at course completion. This is relatively surprising given that our intervention was primarily targeted well before the start of the module, and one might expect that the largest impact would be closest to the initial intervention, in other words, at the start of the course.

Table 5. Chi-square Analysis at Different Course Milestones (Student Module Count)

	Control group		Intervention group		*p* value*
	N=	%	N=	%	
Registrations at course start	279	87.7	294	94.2	.004*
Withdrawn before course start	39	12.3	18	5.8	
Total	318	100%	312	100%	
Registered at 25% FLP	276	86.8	293	93.9	.002*
Withdrawn at 25% FLP	42	13.2	19	6.1	
Total	318	100%	312	100%	
Registered at 50% FLP	234	73.6	275	88.1	.001*
Withdrawn at 50% FLP	84	26.4	37	11.9	
Total	318	100%	312	100%	
Registered at 100% FLP	209	65.7	242	77.6	.001*
Withdrawn at 100% FLP	109	34.3	70	22.4	
Total	318	100%	312	100%	
Reached course completion	146	45.9	175	56.1	.010*
Did not reach course completion	172	54.1	137	43.9	
Total	318	100%	312	100%	

* p =.01

JOURNAL OF LEARNING ANALYTICS

`	Start	FLP1	FLP2	FPL3	Completion
Gross (Control - 318)	279	276	234	209	146
Gross (Intervention - 312)	294	293	275	242	175
Net (uplift)*	21 (7.4%)	23 (8.2%)	46 (19.78%)	38 (18.02%)	32 (22.17%)

*Groups adjusted to size

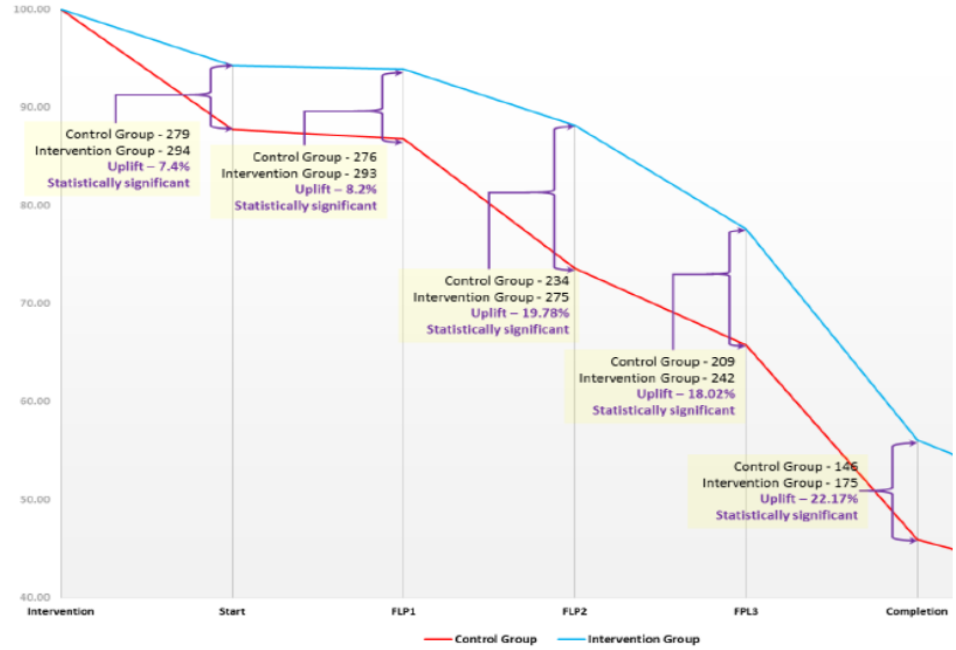

Figure 2. Student retention rates (number of students present) at each course milestone.

In the second step of the analysis, we compared the retention rates of the two different interventions — phone call versus email — to the control group to identify if one of the interventions was more effective. At p<=.01, statistically significant differences were identified at specific milestones, that is, registrations at course start ($\chi2$=8.74, df=2, p=.01), at 25% FLP ($\chi2$=10.05, df=2, p=.001), at 50% FLP ($\chi2$=20.06, df=2, p=.001), and at 100% FLP ($\chi2$=10.77, df=2, p=.005). No statistically significant differences were identified for students reaching course completion ($\chi2$=4.89, df=2, p=.08, NS). Follow-up pairwise comparisons in relation to the four significant milestones above and with an adjusted p value (p=.05/12=.004) showed more student registrations in the phone group at three points compared to the control group: course start, 25% FLP, and 50% FLP. Similarly, the email group performed better compared to the control group at 50% FLP and 100% FLP (see Table 6). Perhaps surprising given previous research (Wilcock, 2011), no statistically significant differences were found between the phone and email groups across milestones. These findings suggest that phoning or emailing students before the start of a course had a similar effect on student retention, yet a single intervention six weeks before the start of the course may not lead to sustained long-term improvements in engagement or course completion. Future studies should examine whether an additional intervention (e.g., closer to the 100% FLP) could further motivate students to complete their studies.

In SPM, predictions or probabilities are based on logistic regression analysis of a number of explanatory variables related to student factors: previous study, students on a course, students' previous progress at the university, and course and qualification variables. This approach of predicting students at risk of not completing their courses could be easily adopted and implemented by other institutions, especially when there are not enough resources to produce more complex statistical models that predict student retention. SPM was shown to be effective in terms of helping SSTs select a specific group of students at risk, based on their probability of completing their studies, and then proactively provide support. This approach of choosing students at risk based on predictive modelling can overcome the potentially limited institutional resource availability, and focus SST efforts on specific cohorts. The study proposes an alternative, effective way of choosing and contacting students at risk that aligns well with and contributes to the evolution of existing practices used by SSTs to identify students at risk. This study suggests that PLA systems such as SPM can be useful tools for different stakeholders, not only teachers (Herodotou et al., 2019a, 2019b) but, also SSTs in distance education for identifying students in most need of assistance and directing available resources to them.

Table 6. Significant Differences Between the Control, Email, and Phone Call Conditions

	Control group		Email group		Phone group		*p* value*
	N=	%	N=	%	N=	%	
Registrations at course start	279	87.7	-	-	127	96.2	.005*
Withdrawn before course start	39	12.3	-	-	5	3.8	
Total	318	100%	-	-	132	100%	
Registered at 25% FLP	276	86.8	-	-	127	96.3	.003*
Withdrawn at 25% FLP	42	13.2	-	-	5	3.8	
Total	318	100%	-	-	132	100%	
Registered at 50% FLP	234	73.6	-	-	117	88.6	.001*
Withdrawn at 50% FLP	84	26.4	-	-	15	11.4	
Total	318	100%	-	-	132	100%	
Registered at 50% FLP	234	73.6	149	87.1	-	-	.001*
Withdrawn at 50% FLP	84	26.4	22	12.9	-	-	
Total	318	100%	171	100%	-	-	
Registered at 100% FLP	209	65.7	136	79.5	-	-	.001*
Withdrawn at 100% FLP	109	34.3	35	20.5	-	-	
Total	318	100%	171	100%	-	-	

* $p<=.004$

In terms of the effectiveness of motivational intervention, outcomes showed that an initial text to students, informing them of an upcoming phone call by SSTs, along with that phone call (or an email for those who can't be reached by phone) asking students about 1) how they feel about starting their course, 2) whether they have any concerns, and 3) whether they know where to look for help, were found to support student retention through to course completion. Our findings indicated that this proactive SST contact with students had positive effects on retention by motivating course progression. These findings align well with existing studies. Early proactive contact via text and phone can reinforce certain positive engagement (Castleman & Page, 2015) while email contact can significantly increase retention (Inkelaar & Simpson, 2015). The timing and type of motivational interventions studied here were found to be effective in terms of encouraging progress and course completion.

These findings could be explained by theories of student persistence suggesting that early involvement with the faculty — during the first 6–7 weeks of semester — is more likely to determine whether students will persist with the institution and succeed in their studies (Milem & Berger, 1997; Lukosius, Pennington, & Olorunniwo, 2013). In this study, the distance learning nature of the university makes it impossible for students to engage with university initiatives and familiarize themselves with faculty life. Hence, interpersonal contact and communication with support and academic teams is more likely to contribute to a sense of belonging and social integration with the university, connecting students with the institution from a distance. Also, the proactive communication of SSTs may have motivated students to seek support and declare any problems and difficulties they face. As a result, students may have received necessary support early on in their studies and resolved any difficulties in a timely manner, without preventing the progress and completion of their studies. Comparing the three conditions — phone call, email, and control groups — it was not clear whether a phone call was more effective than an email, as outcomes showed phone calls to facilitate progression until the 50% FLP while emails were more effective between the 50% and 100% FLPs. This issue requires further examination; in this study only the phone call group communicated with a member of SST to discuss their studies. It is not known whether the email group responded or remained passive.

4. Conclusion

This study described a randomized control trial (RCT) about how predictive learning analytics (PLAs) could proactively be used by university Student Support Teams (SSTs) to identify students at risk and guide targeted motivational interventions. Evidence from this paper suggests that the Student Probabilities Model (SPM) could be used alongside existing motivational interventions to direct resources to students who require special support, and for whom, as data analysis showed, motivational interventions can make a difference in retention. Insights from SPM could inform the teaching strategy in online and distance education. They can be applied to the context of course registration, retention, and completion for identifying students at-risk. They could also be used to identify students with high probabilities of success in order, for example, to encourage early course

ISSN 1929-7750 (online).

enrollment. Motivational interventions such as the targeted use of text, email, or phone can improve administrative support at scale and at low cost.

As an alternative way of choosing and supporting students at risk, predictive modelling aligned well with and contributed to the evolution of existing SST practises used to identify at-risk students at the institution under study. Sclater (2017) highlighted that "Learning analytics requires bringing people with high levels of technical expertise together with others who understand pedagogy and educational processes" (p. 16). Within this study, an interdisciplinary team built on previous and existing student support strategies, and using PLAs, made decisions about identifying students at risk and targeting resource-intensive motivational interventions. It is likely that the close similarity between the proposed motivational intervention and existing practices of identifying and intervening with students more likely favoured successful adoption and implementation by support staff.

In future, information from PLAs could be used to personalize motivational interventions by, for example, tailoring contact depending on a student's probability band. Thus far, SPM has been used to identify students at risk and adapt the SST intervention strategy. Predictive data can also be used to personalize the content of the motivational interventions. For example, one of the current SST interventions at the university under study refers to "a follow up for assignment non-submission," which is highly recommended for all non-submitting students without a recorded extension.

A challenge yet to be addressed is the alignment between contact made by SSTs and interventions made by teachers. A system that records actions made by both "actors" to ensure that overlaps are avoided and students receive appropriate support is still needed. At the institution under study, teachers have been given access to a predictive dashboard combining data from different tools, including SPM, to better support their practice. The likelihood of a student being at risk is expected to inform teacher practices, hopefully leading to effective, proactive student support.

Declaration of Conflicting Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors declared no financial support for the research, authorship, and/or publication of this article.

References

Arnold, K., & Pistilli, M. (2012). Course Signals at Purdue: Using learning analytics to increase student success. In S. Buckingham Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (LAK '12), 29 April–2 May 2012, Vancouver, BC, Canada (pp. 267–270). New York: ACM.

Beard, J., Stevens, C., & Welch, A. (2017). *Undergraduate MILLS framework 2017/18. Student Support Centre for Excellence, Open University, UK.*

Calvert, C. (2014). Developing a model and applications for probabilities of student success: A case study of predictive analytics. *Open Learning: The Journal of Open, Distance and e-Learning*, *29*(2), 160–173. https://dx.doi.org/10.1080/02680513.2014.931805

Castleman, B. L., & Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior and Organization*, *115*, 144–160. https://dx.doi.org/10.1016/j.jebo.2014.12.008

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, *10*(1), 17–29. https://dx.doi.org/10.1109/tlt.2016.2616312

Dawson, S., Jovanovic, J., Gašević, D., & Pardo, A. (2017). From prediction to impact: Evaluation of a learning analytics retention program. *Proceedings of the 7th International Conference on Learning Analytics and Knowledge* (LAK '17), 13–17 March 2017, Vancouver, BC, Canada (pp. 474–478). New York: ACM.

Dolata, U. (2013). *The transformative capacity of new technologies: A theory of sociotechnical change*. London: Routledge.

Ferguson, R., Brasher, A., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., ... Vuorikari, R. (2016). A European framework for action on learning analytics. In R. Vuorikari & J. Castaño-Muñoz (Eds.), *Research evidence of the use of learning analytics: Implications for education policy* (pp. 1–152). Luxembourg: Publications Office of the European Union.

Ferguson, R., & Clow, D. (2017). Where is the evidence? *Proceedings of the 7th International Conference on Learning Analytics and Knowledge* (LAK '17), 13–17 March 2017, Vancouver, BC, Canada (pp. 56–65). New York: ACM.

Fozdar, B., & Kumar, L. (2007). Mobile learning and student retention. *International Review of Research in Open and Distance Learning*, *8*(2), 1–18. https://dx.doi.org/10.4103/0973-029X.102522

Harley, J. M., Carter, C. K., Papaionnou, N., Bouchet, F., Landis, R. S., Azevedo, R., & Karabachian, L. (2016). Examining the predictive relationship between personality and emotion traits and students' agent-directed emotions: Towards emotionally-adaptive agent-based learning environments. *User Modeling and User-Adapted Interaction*, *26*(2–3), 177–219. https://dx.doi.org/10.1007/s11257-016-9169-7

Herodotou, C., Rienties, B., Hlosta, M., Boroowa, A., Mangafa, C., & Zdrahal, Z. (2020). The scalable implementation of predictive learning analytics at a distance learning university: Insights from a longitudinal case study. *The Internet and Higher Education*, *45*. https://dx.doi.org/10.1016/j.iheduc.2020.100725

Herodotou, C., Hlosta, M., Boroowa, A., Rienties, B., Zdrahal, Z., & Mangafa, C. (2019a). Empowering online teachers through predictive learning analytics. *British Journal of Educational Technology*, *50*(6), 3064–3079. https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12853

Herodotou, C., Rienties, B., Boroowa, A., Zdrahal, Z., & Hlosta, M. (2019b). A large-scale implementation of predictive learning analytics in higher education: The teachers' role and perspective. *Educational Technology Research and Development*, *67*(5), 1273–1306. https://dx.doi.org/10.1007/s11423-019-09685-0

Hlosta, M., Zdrahal, Z., & Zendulka, J. (2017). Ouroboros: Early identification of at-risk students without models based on legacy data. *Proceedings of the 7th International Conference on Learning Analytics and Knowledge* (LAK '17), 13–17 March 2017, Vancouver, BC, Canada (pp. 6–15). New York: ACM.

Hung, I. C., Chao, K. J., Lee, L., & Chen, N. S. (2013). Designing a robot teaching assistant for enhancing and sustaining learning motivation. *Interactive Learning Environments*, *21*(2), 156–171. https://dx.doi.org/10.1080/10494820.2012.705855

Inkelaar, T., & Simpson, O. (2015). Challenging the "distance education deficit" through "motivational emails." *Open Learning*, *30*(2), 152–163. https://dx.doi.org/10.1080/02680513.2015.1055718

Jayaprakash, S. M., Moody, E. W., Lauría, E. J. M., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, *1*(1), 6–47. https://dx.doi.org/10.18608/jla.2014.11.3

Kim, C. (2008). Using email to enable e3 (effective, efficient, and engaging) learning. *Distance Education*, *29*(2), 187–198. https://dx.doi.org/10.1080/01587910802154988

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*, 18–33. https://dx.doi.org/10.1016/j.compedu.2016.10.001

Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). Penetrating the black box of time-on-task estimation. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge* (LAK '15), 16–20 March 2015, Poughkeepsie, NY, USA (pp. 184–193). New York: ACM.

Lukosius, V., Pennington, J. B., & Olorunniwo, F. O. (2013). How students' perceptions of support systems affect their intentions to drop out or transfer out of college. *Review of Higher Education & Self-Learning*, *6*(18), 209–221.

Milem, J. F., & Berger, J. B. (1997). A modified model of college student persistence: Exploring the relationship between Astin's theory of involvement and Tinto's theory of student departure. *Journal of College Student Development*, *38*(4), 387–400.

Moore, M. G. (1997). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical principles of distance education* (pp. 22–38). London: Routledge.

Naismith, L. (2007). Using text messaging to support administrative communication in higher education. *Active Learning in Higher Education*, *8*(2), 155–171. https://dx.doi.org/10.1177/1469787407078000

Ryan, G. S., & Davies, F. (2016). Development and evaluation of an online, interactive information and advice tool for pre-registration nursing students. *Nurse Education in Practice*, *17*, 214–222. https://dx.doi.org/10.1016/j.nepr.2016.01.002

Scheffel, M., Drachsler, H., de Kraker, J., Kreijns, K., Slootmaker, A., & Specht, M. (2017). Widget, widget on the wall, am I performing well at all? *IEEE Transactions on Learning Technologies*, *10*(1), 42–52. https://dx.doi.org/10.1109/tlt.2016.2622268

Sclater, N. (2017). *Learning analytics explained*. New York: Routledge.

Sclater, N., Peasgood, A., & Mullan, J. (2016). Learning analytics in higher education: A review of UK and international practice. Retrieved from https://go.aws/328xLVv

Seidel, E., & Kutieleh, S. (2017). Using predictive analytics to target and improve first-year student attrition. *Australian Journal of Education*, *61*(2), 200–218. https://dx.doi.org/10.1177/0004944117712310

Shea, P., & Bidjerano, T. (2014). Does online learning impede degree completion? A national study of community college students. *Computers & Education, 75*(Supplement C), 103–111. https://dx.doi.org/10.1016/j.compedu.2014.02.009

Simpson, O. (2013). Student retention in distance education: Are we failing our students? *Open Learning, 28*(2), 105–119. https://dx.doi.org/10.1080/02680513.2013.847363

Tempelaar, D., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018). Student profiling in a dispositional learning analytics application using formative assessment. *Computers in Human Behavior*, *78*, 408–420. https://dx.doi.org/10.1016/j.chb.2017.08.010

van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers & Education*, *90*, 80–94. https://dx.doi.org/10.1016/j.compedu.2015.09.006

van Ameijde, J., Weller, M., & Cross, S. (2018). Learning design for student retention. *Journal of Perspectives in Applied Academic Practice, 6*(2), 41–50. https://dx.doi.org/10.14297/jpaap.v6i2.318

Wentzel, K. R., & Wigfield, A. (2007). *Motivational interventions that work: Themes and remaining issues, 1520*, 261–271. https://dx.doi.org/10.1080/00461520701621103

West, D., Luzeckyj, A., Searle, B., Toohey, D., & Price, R. (2018). *The use of learning analytics to support improvements in teaching practice.* Innovative Research Universities. Melbourne, Australia.

Wilcock, C. (2011). The Open University in Wales: Proactive student support. In L. Thomas & C. Jamieson-Bal (Eds.), *Engaging students to improve student retention and success in higher education in Wales*. York, UK: Higher Education Academy.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. *Proceedings of the 4th International Conference on Learning Analytics and Knowledge* (LAK '14), 24–28 March 2014, Indianapolis, IN, USA (pp. 203–211). New York: ACM.

Worsley, M., & Blikstein, P. (2015). Leveraging multimodal learning analytics to differentiate student learning strategies. *Proceedings of the 5th International Conference on Learning Analytics and Knowledge* (LAK '15), 16–20 March 2015, Poughkeepsie, NY, USA (pp. 360–367). New York: ACM.

Wright, M. C., McKay, T., Hershock, C., Miller, K., & Tritz, J. (2014). Better than expected: Using learning analytics to promote student success in gateway science. *Change: The Magazine of Higher Learning*, *46*(1), 28–34. https://dx.doi.org/10.1080/00091383.2014.867209